THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 266, No. 4, Issue of February 5, pp. 2297-2302, 1991
Printed in U.S.A.

EXHIBIT G

Phosphorylation Causes a Conformational Change in the Carboxyl-terminal Domain of the Mouse RNA Polymerase II Largest Subunit*

(Received for publication, August 3, 1990)

Jie Zhang and Jeffry L. Corden‡

From the Howard Hughes Medical Institute and Department of Molecular Biology and Genetics, The Johns Hopkins University School of Medicine, Baltimore, Maryland 21205

The carboxyl-terminal domain (CTD) of the largest subunit of eukaryotic RNA polymerase II can be phosphorylated by a $p34^{cdc2/CDC28}$-containing CTD kinase. Phosphorylated serine (or threonine) is located at positions 2 and 5 in the repetative heptapeptide consensus sequence Tyr^1-Ser^2-Pro^3-Thr^4-Ser^5-Pro^6-Ser^7. We show here that phosphorylation of the mouse CTD retards its electrophoretic mobility in sodium dodecyl sulfate-polyacrylamide gels in a way similar to that observed for the II_o form of the largest subunit of RNA polymerase II phosphorylated *in vivo*. At the maximum level of phosphorylation by CTD kinase *in vitro*, there are 15–20 phosphates evenly distributed among the 52 heptapeptide repeats that comprise the mouse CTD. Gel filtration chromatography and sucrose gradient ultracentrifugation analyses indicate that phosphorylation induces a dramatic conformational change in the CTD with the phosphorylated form adopting a far more extended structure than the unphosphorylated CTD.

RNA polymerase II contains an unusual extension of its largest subunit that is not present in other multisubunit RNA polymerases. This carboxyl-terminal domain (CTD)[1] is comprised of tandem repeats of the consensus sequence Tyr^1-Ser^2-Pro^3-Thr^4-Ser^5-Pro^6-Ser^7 (for reviews see Corden, 1990; Woychik and Young, 1990). The heptapeptide sequence is present in RNA polymerase II from a variety of species, being repeated from 17 times in *Plasmodium falciparum* (Li *et al.*, 1989) to 52 times in mammals (Corden *et al.*, 1985). Genetic studies have shown that this domain plays an essential role in RNA polymerase II function (Allison *et al.*, 1988; Bartolomei *et al.*, 1988; Nonet *et al.*, 1987; Zehring *et al.*, 1988); however, the nature of this essential function is unknown.

The CTD can be hyperphosphorylated *in vivo*, and this phosphorylation is associated with an increase in the apparent molecular mass of the largest subunit of >20 kDa, greater than what can be accounted for by an increase in mass of added phosphate (Cadena and Dahmus, 1987). SDS-gel electrophoresis has demonstrated that the phosphorylated form of the largest subunit is the predominant form *in vivo* (Kim and Dahmus, 1986) suggesting that phosphorylation has an important function. Recent experiments have implicated CTD phosphorylation in the shift from transcription initiation to elongation (Laybourn and Dahmus, 1989, 1990).

Protein kinases that phosphorylate the CTD have been isolated either on the basis of their ability to phosphorylate heptapeptide substrates (Cisek and Corden, 1989; Stevens and Maupin, 1989) or on their ability to cause the characteristic shift in electrophoretic mobility of the IIa subunit (Guilfoyle, 1989; Lee and Greenleaf, 1989; Payne *et al.*, 1989). Here we show that a mouse CTD kinase can cause a shift in the electrophoretic mobility of a mouse CTD substrate. Stoichiometric labeling results in 15–20 phosphates distributed evenly among the 52 repeats in the mouse CTD and hydrodynamic studies show that this phosphorylation results in a dramatic conformation change in the CTD.

EXPERIMENTAL PROCEDURES

In Vitro Phosphorylation of CTDα—For analysis of the time course of CTD phosphorylation, a 50-μl reaction was set up in kinase buffer (60 mM KCl, 50 mM Tris-HCl, pH 7.8, 10 mM $MgCl_2$, 0.1 mM dithiothreitol) with 100 pmol of CTD (prepared as described in the accompanying paper), 100 ng of CTD kinase *E*2 (Cisek and Corden, 1990) and 4 mM [γ-^{32}P]ATP at a specific radioactivity of 570 cpm/pmol. The reaction mixture was kept at 30 °C and 5-μl aliquots were taken at specific time points. The aliquots were mixed with 5 μl of 2 × SDS sample buffer containing 20 mM EDTA, boiled at 100 °C for 3 min and loaded on 10% SDS gel. After electrophoresis the SDS gel was vacuum dried on 3MM paper and exposed to x-ray film.

Determination of the Stoichiometry of Phosphorylated CTDα—The extent of CTD phosphorylation was determined in two separate experiments. In the first experiment a known amount of CTD substrate was labeled with [γ-^{32}P]ATP of a known specific radioactivity by CTD kinase *E*2. After SDS-PAGE the ^{32}P-labeled CTD was excised from the gel using the autoradiogram as a template, and its radioactivity was determined by liquid scintillation counting. The amount of phosphate on the CTD was calculated from the specific radioactivity of [γ-^{32}P]ATP and the number of phosphates/CTD was obtained from the ratio of the amount of phosphate over the amount of CTD in the aliquot.

In a second experiment, 240 pmol of CTD and 100 ng of CTD kinase *E*2 were incubated at 30 °C for 16 h in 40 μl of kinase buffer with [γ-^{32}P]ATP at a specific radioactivity of 7.25 cpm/pmol. The reaction mixture was then chromatographed on an FPLC Superose 6 HR 10/30 column (see below) and 400-μl fractions were collected. Fractions containing fully phosphorylated ^{32}P-CTD were identified by SDS-PAGE, and the concentration of the CTD was determined from its absorbance at 274 nm (assuming the extinction coefficient for the CTD at 274 nm was $52 \times 1.4 \times 10^3$ cm^{-1} M^{-1}, 52 times the molar extinction coefficient for tyrosine). 200 μl of this [^{32}P]CTD was mixed with 10 ml of high flash point mixture (Research Product International Corp.), and the radioactivity was determined by scintillation counting. The amount of phosphate on ^{32}P-CTD was calculated from its radioactivity divided by the specific radioactivity of [γ-^{32}P]ATP.

The [γ-^{32}P]ATP used in the above experiments was prepared by mixing the appropriate amount of [γ-^{32}P]ATP (Du Pont-New England Nuclear) with preweighed ATP (Sigma) that was freshly dissolved in deionized distilled water and adjusted to pH 7.0 with 1 N NaOH. A 40-μl aliquot of ATP was loaded on an FPLC Mono Q HR 5/5 column equilibrated with 1% buffer B (40 mM Tris-HCl, pH 8.0,

* The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

‡ To whom correspondence should be addressed.

[1] The abbreviations used are: CTD, carboxyl-terminal domain; SDS-PAGE, sodium dodecyl sulfate-polyacrylamide gel electrophoresis; FPLC, fast protein liquid chromatography.

2 M NaCl) and eluted by a linear gradient with an increasing of buffer B (8%/ml) to 100%. The ultraviolet spectrum of the most radioactive fraction was measured to obtain an A_{259} value for concentration determination. The extinction coefficient for ATP at 259 nm was taken as 1.54×10^4 cm^{-1} M^{-1}. An aliquot of this ATP fraction was mixed with 10 ml of high flash point mixture, counted in a liquid scintillation counter, and the specific activity (cpm/pmol) of [γ-^{32}P] ATP was determined by calculating the ratio of radioactivity over the amount of ATP.

Distribution of Phosphates on the CTD—Continuous ^{32}P-labeling of the CTD by CTD kinase *E*2 was as described in the accompanying paper. At the end of phosphorylation either 1 μg of sequencing grade trypsin (Boehringer Mannheim) or 1 μg of *Staphylococcus aureus* (V8) protease (Pierce Chemical Co.) was added to the 10-μl reaction for protease cleavage at 37 °C (trypsin) or 30 °C (V8 protease) for 1 h. After protease digestion, 10 μl of 2 × SDS sample buffer was added, and the sample was electrophoresed on a 12.5% SDS gel. Determination of radioactivities of the ^{32}P-labeled CTD protease fragments was as described above for that of ^{32}P-CTD. Relative percentages of phosphates on the protease fragments were calculated with the complete ^{32}P-CTD as 100%.

Pulse-chase ^{32}P-labeling was the same as above except that the initial specific radioactivity for [γ-^{32}P]ATP was 3250 Ci/mmol (28 μM). After 30 min the ATP concentration was adjusted to 4 mM with non-radioactive ATP, and the reaction was kept at 30 °C for another 3 h. The rest of the analysis was as described above for that of continuously labeled ^{32}P-CTD protease fragments.

Phosphorylation of the CTD by Casein Kinase II—Phosphorylation of the CTD by casein kinase II was the same as described for CTD kinase but with 5 ng of casein kinase II (purified by Dr. Lars Cisek (Johns Hopkins Medical School) according to the method of Zandomeni *et al.*, 1986) replacing CTD kinase. The ^{32}P-CTDa phosphorylated by casein kinase II and its protease fragments were analyzed as described above for CTD kinase-labeled ^{32}P-CTD_0.

Urea-PAGE Separation of Phosphorylated CTD Intermediates—20 pmol of CTD was phosphorylated as described above by casein kinase II in a final volume of 10 μl with [γ-^{32}P]ATP at a final concentration of 0.1 mM and a specific radioactivity of 1000 Ci/mmol. After 1 h of incubation at 30 °C, the volume of the reaction mixture was adjusted to 50 μl with kinase buffer, the ATP concentration was adjusted to 4 mM with nonradioactive ATP, and 100 ng of CTD kinase was added. At specific time points, 5-μl aliquots were taken and immediately boiled in 5 μl of sample buffer with 2 M urea substituting for SDS. Discontinuous urea-polyacrylamide gels with 5% resolving gels and 4% stacking gels were prepared as for conventional SDS gels (Laemmli, 1970) except that 2 M urea replaced the SDS. SDS was also omitted from the running buffer, and these gels were run for 1.5 times the time taken for the bromphenol blue dye to run off the gel.

Gel Filtration Chromatography of Phosphorylated and Unphosphorylated CTD—Aliquots of ^{32}P-labeled CTD, either phosphorylated by CTD kinase *E*2 or phosphorylated by casein kinase II were loaded onto a Superose 6 HR 10/30 column equilibrated with kinase buffer, and the proteins were eluted in the same buffer. Aliquots from each of the 400-μl fractions were electrophoresed on 10% SDS gels and CTD-containing fractions were identified by autoradiography. Standards used to calibrate the column were ferritin, catalase, aldolase, bovine serum albumin, enolase, and ovalbumin, whose Stoke's radii in units of 10^{-8} cm are 79.0, 52.2, 47.4, 37.0, 34.1, and 27.6, respectively.

Sucrose Gradient Centrifugation of Phosphorylated and Unphosphorylated CTD—10 pmol of ^{32}P-CTD_0 phosphorylated by CTD kinase *E*2 was mixed with 10 μg each of sedimentation standards: aldolase (7.35 S, 1 S = 10^{-13} s), ovalbumin (3.55 S) and ribonuclease (2.00 S) in a final volume of 100 μl kinase buffer. 10 pmol of ^{32}P-CTDa phosphorylated by casein kinase II was mixed with 10 μg each of sedimentation standards: bovine serum albumin (4.31 S), carbonic anhydrase (2.85 S) and cytochrome c (1.83 S) in a final volume of 100 μl of kinase buffer. The above samples were layered onto two 5.1-ml linear 5–18% sucrose gradients prepared in kinase buffer and ultracentrifuged at 45,000 rpm in a Beckman SW 50.1 rotor for 24 h at 20 °C. After centrifugation, 350-μl fractions were collected dropwise by puncturing the bottoms of the tubes and 30-μl aliquots were run on 15% SDS gels. After electrophoresis the gel was silver stained according to a method of Morrisey (1981) to reveal the standards. The gel was then dried and exposed to x-ray film to detect the CTD-containing fractions by autoradiography.

RESULTS

Phosphorylation of CTDa Causes a Shift of Its Electrophoretic Mobility—The use of a complete mouse CTD substrate has allowed us to carefully examine the phosphorylation reaction catalyzed by mouse CTD kinase *E*2. In the preceding paper we showed that phosphorylation of the CTD by mouse CTD kinase *E*2 retards its mobility in SDS gels to a position similar to that seen when the *in vivo* phosphorylated CTD is excised from subunit II_0 by cyanogen bromide cleavage.

In this paper we demonstrate the presence of intermediates in the phosphorylation reaction and show that during the course of the phosphorylation reaction two distinct phases are detected by SDS-gel electrophoresis. During the early stages of the reaction the mobility of the CTD is progressively retarded, while later in the reaction the mobility of the CTD actually increases with addition of more phosphates (Fig. 1*a*). No discrete intermediate bands are detected but rather the intermediates form a continuum. We interpret the two stages of this reaction in the following way. The earlier phosphates added preclude the binding of a number of SDS molecules thus decreasing the charge to mass ratio of the CTD. During this stage the more phosphates added, the more the mobility is retarded. When about 40% of the maximum number of phosphates have been added, no SDS can bind to the CTD, and the addition of further phosphates then increases the charge to mass ratio and hence the electrophoretic mobility.

Fig. 1*b* shows the result of a similar CTD kinase reaction. In this case, however, the CTD substrate has been prelabeled by phosphorylation with casein kinase II which recognizes only 1 or 2 serine residues at the extreme carboxyl terminus (see "Experimental Procedures" and Fig. 3). An excess of nonradioactive ATP is added to the reaction before CTD kinase *E*2 is added, and the resulting reaction products, taken at various times, are resolved by urea-PAGE electrophoresis (see "Experimental Procedures"). Under these conditions, we see that the addition of phosphates to the CTD substrate increases the charge to mass ratio of the CTD and thus increases its electrophoretic mobility. This reaction is essentially completed at 160 min. Unlike the SDS gel system, the urea gel system does not demonstrate a biphasic mobility shift. The urea-PAGE system is also able to resolve the partially phosphorylated intermediates generated toward the end of the reaction; intermediates are not well resolved on SDS gels.

CTD Kinase E2 Adds Less Than One Phosphate/Heptapeptide Repeat—We have shown by sequencing that CTD kinase *E*2 transfers phosphates to serines and threonines that precede proline at positions 2 and 5 in the heptapeptide repeat (see accompanying paper). In the mouse CTD there are 95 Ser-Pro and 9 Thr-Pro-recognition sites for the mouse CTD kinase *E*2, and we have determined the number of these sites phosphorylated in CTD_0 either by excising the labeled band from an SDS gel and measuring its radioactivity in a scintillation counter or by purifying labeled CTD by gel filtration chromatography and determining its concentration by absorbance spectroscopy and its radioactivity by scintillation counting. Because we know the specific activity of the [γ-^{32}P] ATP and the concentration of the CTD substrate (see "Experimental Procedures") we can calculate the number of phosphates present on the CTD at the end of the reaction. Calculated from these data, there are at maximum 15–20 phosphates distributed among the 52 heptapeptide repeats of the mouse CTD (Fig. 1*c*).

Distribution of Phosphates on CTD_0—CTD kinase *E*2 phosphorylates only 15–20 of 104 possible sites in the mouse CTD. Are the phosphorylated residues randomly distributed among

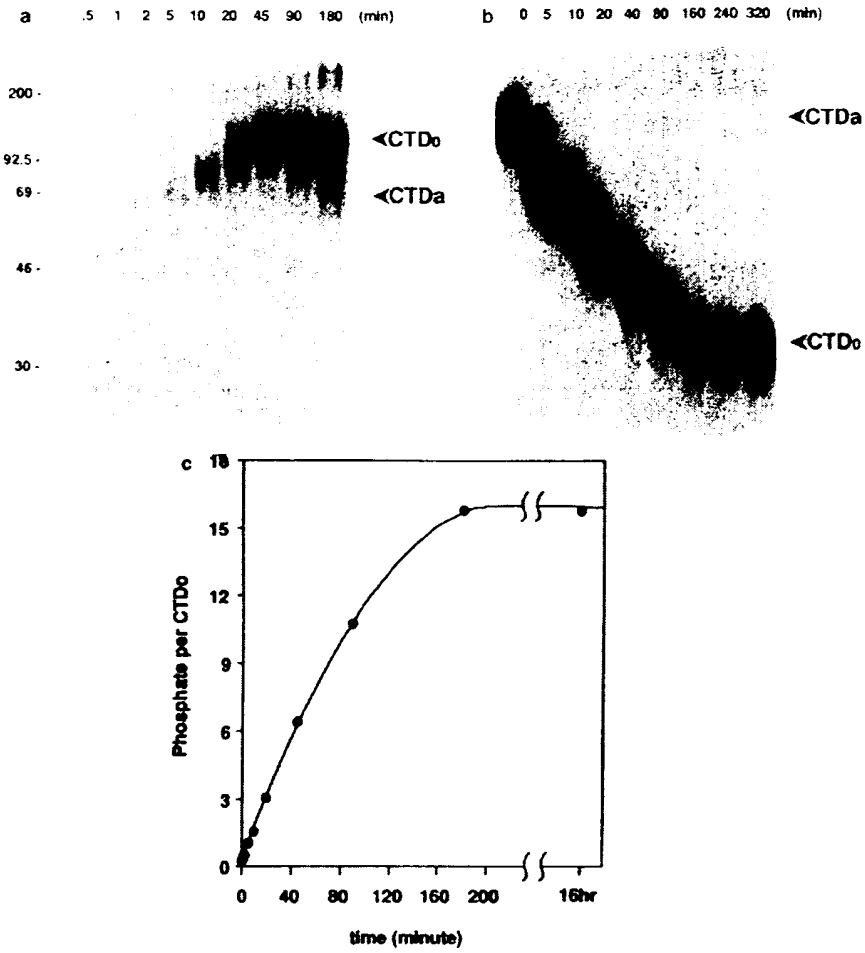

FIG. 1. **Time course of *in vitro* phosphorylation of the CTD by CTD kinase *E*2.** *Panel a* shows a continuous labeling of CTD by CTD kinase *E*2 in the presence of [γ-^{32}P]ATP. At the time points indicated, aliquots were taken and reactions were terminated by SDS sample buffer containing 20 mM EDTA. Samples were separated on a 10% SDS gel that was dried for autoradiography after electrophoresis. The standard molecular masses are indicated in kDa at the left and the positions of unphosphorylated CTD (*CTDa*) and completely phosphorylated CTD (CTD_o) at right. *Panel b* shows a pulse-chase labeling of CTD. Initial labeling of CTD was carried out by casein kinase II at high specific radioactivity and low concentration of [γ-^{32}P]ATP. Further phosphorylation was achieved by adding CTD kinase at time zero after increasing the nonradioactive ATP concentration. The phosphorylation time course was followed by taking aliquots at the specific time points and separating CTD intermediates on urea-PAGE (see "Experimental Procedures"). The gel was dried and exposed to show the autoradiograph. *Panel c*, the number of phosphates/molecule of CTD at each time point during continuous labeling (in *panel a*) was determined by scintillation counting of the ^{32}P-CTD regions, calculated from the known specific radioactivity of [γ-^{32}P]ATP and the known concentration of CTD, and plotted as a function of time.

the 52 heptapeptide repeats or are they clustered? We have taken advantage of several naturally occurring protease sites in the mouse CTD to generate CTD fragments and have examined their degree of phosphorylation. *S. aureus* V8 protease cuts the mouse CTD into two fragments, a large one, from repeat 2 to 37, and a small one from 37 to 52. Trypsin digestion generates a fragment encompassing the region from repeat 2 to 31 and a number of small fragments ranging from one to four repeats (Fig. 2*a*). Both continuously and pulse-chase CTD kinase *E*2-labeled ^{32}P-CTD were cleaved by trypsin or V8 protease, and the proteolytic fragments of the CTD were resolved on a 12.5% SDS gel (Fig. 2*b*). All of the three major fragments, the largest tryptic fragment and the large and small V8 fragments, are phosphorylated, indicating phosphates are not clustered. Quantification of the relative distribution of phosphates was determined by scintillation counting of the excised bands, and the results are summarized in Table I. In continuous labeling, the percentage of phosphates on each of the fragments is close to the percentage of the heptapeptide repeats in the fragment, indicating a nearly even distribution of phosphates on CTD. The phosphate percentage of the small V8 fragment is slightly higher than its repeat percentage, while both V8 large and tryptic fragments have phosphate percentages lower than their repeat percentages. This suggests that the non-consensus region at the end of the CTD may actually have on average a few more phosphates than the consensus region (see Fig. 5 in Corden *et al.*, 1985 for complete CTD sequence). The distribution of label in the pulse-labeled sample is more non-random with a higher percentage of phosphate in the more carboxyl-terminal repeats. These distal repeats contain a number of Ser^5-Pro^6-Lys^7 sequences (as compared with the Ser^5-Pro^6-Ser^7 in the consensus repeats). The basic regions are closer to the emerging p34$^{cdc2/CDC28}$ [2] recognition site and have been shown to be preferentially phosphorylated by CTD kinase *E*2 (see accompanying paper).

Phosphorylation of CTDa by Casein Kinase II—Dahmus (1981) showed that casein kinase II phosphorylates subunit IIa *in vitro*, and the site of this phosphorylation has recently been localized to the CTD (Payne *et al.*, 1989). When our purified CTDa is ^{32}P-labeled by casein kinase II, it migrates on SDS gels with an electrophoretic mobility indistinguishable from that of unphosphorylated CTDa detected by immunoblot (Fig. 3, *lane 1*). After cleavage of casein kinase II-labeled ^{32}P-CTDa, phosphates are associated with the smaller fragment after V8 protease digestion (Fig. 3, *lane 3*) and with small (1–4 heptapeptide repeat) tryptic peptides (Fig. 3, *lane 2*). Taken together with our previous results that synthetic CTD heptapeptides are poor substrates for casein kinase II (Cisek and Corden, 1989), these results suggest that the site recognized by casein kinase II lies outside the heptapeptide repeat region, possibly in the carboxyl-terminal sequence

[2] p34$^{cdc2/CDC28}$ is the catalytic subunit of a class of cell-cycle regulated kinases. This 34-kDa subunit is encoded by *cdc2*$^+$ gene in *Schizosaccharomyces pombe*, and *CDC28* in *Saccharomyces cerevisiae*. See Nurse, 1990 for review.

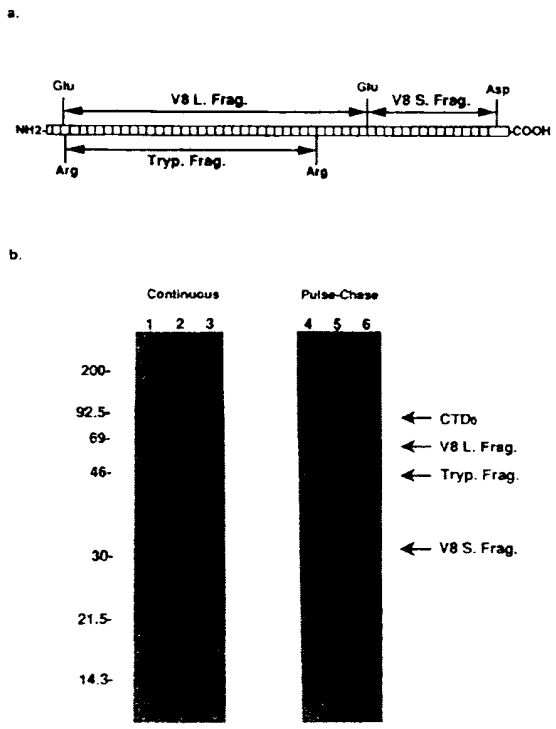

FIG. 2. **Distribution patterns of phosphates on CTD.** *a*, schematic map of the CTD is shown to illustrate the protease cleavage sites in mouse CTD. Each *box* represents one heptapeptide repeat. *S. aureus* V8 protease cuts at the carboxyl sides of glutamic acids in repeats 2, 37, and 52, generating two major fragments, V8 large and small. Trypsin cuts at the carboxyl sides of arginines in repeats 2 and 31, as well as of 8 lysines in repeats 35, 38, 39, 40, 42, 45, 47, and 49 (not shown), producing a major tryptic fragment and several small peptides. *b*, ^{32}P-CTD from continuous (*lanes 1–3*) or pulse-chase labelings (*lanes 4–6*) were loaded either directly or after protease cleavages onto a 12.5% SDS gel. The autoradiogram is shown here with molecular mass values indicated in kDa at the left and positions of uncut CTD (for *lane 1* and *4*), tryptic fragment (for *lanes 2* and *5*), and V8 large and small fragments (for *lanes 3* and *6*) at the right. See Fig. 5 in Corden *et al.*, 1985 for complete sequence of CTD.

TABLE I

Relative phosphate contents of CTD$_0$ fragments

	No. of Repeats	Percentage of Repeats of CTD$_0$	Percentage of Phosphates	
			Continuous	Pulse-chase
CTD$_0$	52.5	100.0	100.0	100.0
V8 large fragment	35.5	67.6	55.4	42.3
Tryptic fragment	29.0	56.2	51.1	35.9
V8 small fragment	15.0	28.6	30.5	39.4

AISPDDSDEEN-COOH. Casein kinase II thus provides a sensitive method for labeling the CTDa without shifting its electrophoretic mobility (Payne *et al.*, 1989).

A Conformational Change Associated with the Mobility Shift of Phosphorylated CTD—The primary means of demonstrating the shift from the IIa to the II$_0$ form of the CTD has been SDS-gel electrophoresis. The results of this paper suggest that a change in SDS binding is one cause of this shift in mobility. It has been proposed, based on an increase in protease sensitivity of the CTD, that phosphorylation also causes a conformational change (Laybourn and Dahmus, 1989). Here we show directly that this is the case.

CTD kinase *E*2-labeled ^{32}P-CTD$_0$ and casein kinase II-labeled ^{32}P-CTDa were centrifuged at 45,000 rpm for 24 h through a 5–18% sucrose gradient together with a set of

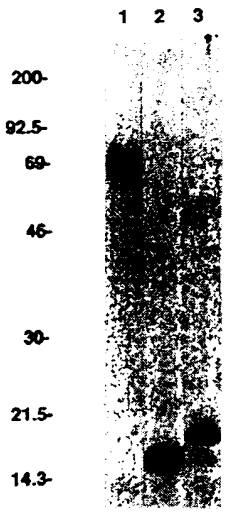

FIG. 3. ***In vitro* phosphorylation of CTD by casein kinase II.** CTDa was incubated with mouse casein kinase II in the presence of [γ-^{32}P]ATP and the ^{32}P-CTDa was loaded either directly (*lane 1*) or after trypsin (*lane 2*) and V8 protease (*lane 3*) digestions onto a 12.5% SDS gel. See Fig. 2*a* for the cleavage sites of trypsin and V8 protease on CTD. The autoradiogram of the gel is shown with molecular mass values indicated in kDa at the left.

standard proteins and analyzed by SDS-PAGE (see "Experimental Procedures"). Svedberg values ($S_{20,w}$) for CTD$_0$ and CTDa obtained from the standard curve are 2.9 S and 2.2 S, respectively (Fig. 4*a*). Such a major difference in sedimentation coefficients cannot be explained by the mass difference between CTDa and CTD$_0$, which is only 3% assuming 15 phosphates. Assuming that the CTD approximates a free-draining polymer, CTD$_0$ and CTDa would be expected to have sedimentation coefficients independent of their masses (Cantor and Schimmel, 1980). In this case, the main cause of different sedimentation behaviors for CTD$_0$ and CTDa would derive from differences in their degree of extension.

Significant conformational differences between CTDa and CTD$_0$ were also apparent in gel filtration chromatography. From the standard calibration curve derived from Superose 6 chromatography, CTDa has an apparent Stoke's radius of 3.8 $\times 10^{-7}$ cm while CTD$_0$ has a value of 8.2 $\times 10^{-7}$ cm (Fig. 4*b*). CTDa with a M_r of 39,883 elutes close to bovine serum albumin, a 65-kDa protein, indicating that CTDa has a slightly extended structure. CTD$_0$ elutes much earlier, at a point close to ferritin, a 130-kDa protein, indicating that CTD$_0$ is highly extended. Partially phosphorylated CTDs elute at points between CTDa and CTD$_0$ (data not shown), suggesting that the changes are continuous rather than cooperative. The apparent Stoke's radii of CTD$_0$ and CTDa are independent of concentration over a 100-fold range, indicating that aggregation is not occurring. The results of these hydrodynamic studies thus indicate that the electrophoretic mobility difference between CTD$_0$ and CTDa corresponds to a conformation change caused by phosphorylation.

DISCUSSION

We have previously described the identification and purification of CTD kinases from mouse ascites tumor cells (Cisek and Corden, 1989, 1990). The CTD kinase used in this study, CTD kinase *E*2, contains a p34$^{cdc2/cdc28\ 2}$ catalytic subunit and recognizes Ser-Pro and Thr-Pro sites in the CTD (see accompanying paper). The ratio of phosphoserine to phosphothreonine in the CTD labeled *in vitro* by CTD kinase *E*2 is similar

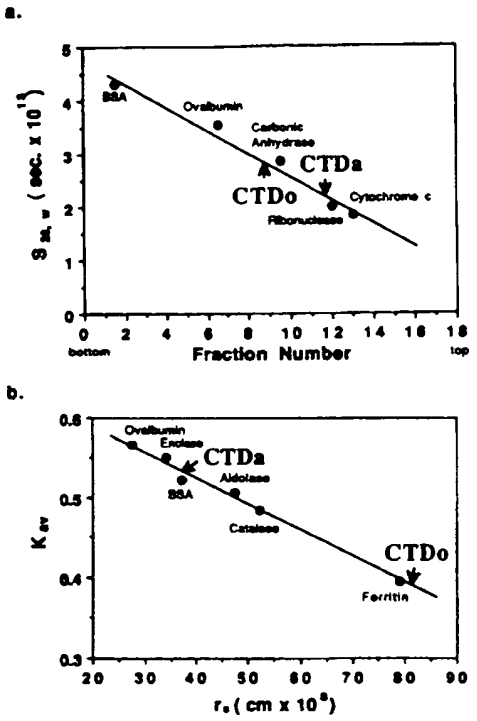

FIG. 4. **Hydrodynamic properties of the phosphorylated and unphosphorylated CTD.** *a*, determination of sedimentation coefficients for CTD_0 and CTDa. ^{32}P-Labeled CTD_0 and CTDa were layered on linear 5–18% sucrose gradients with standards and centrifuged at 45,000 rpm × 24 h. Aliquots of the fractions were run on a 15% SDS gel. Standards in the fractions were detected by silver staining, and CTD_0 and CTDa were revealed by autoradiography. The sedimentation coefficients of standards are plotted as a function of fraction number. The positions of CTD_0 and CTDa are indicated with *arrows* and correspond to Svedberg values of 2.91 S and 2.24 S, respectively. *b*, determination of the apparent Stoke's radii of CTD_0 and CTDa. ^{32}P-Labeled CTD_0 and CTDa were chromatographed by gel filtration on an FPLC Superose 6 HR 10/30 column. Elution points for CTD_0 and CTDa were determined by autoradiograph after SDS-PAGE. Elution volumes for standards were determined by their UV absorbances at 280 nm on the FPLC UV monitor. For each species the partition coefficient K_{av} was calculated from its elution volume (V_e), the void volume of the column (V_0 = 7.85 ml, as determined by the elution point of dextran blue) and the total volume of the column (V_t = 15.7 ml), according to the equation $K_{av} = (V_e - V_0)/(V_t - V_0)$. K_{av} is plotted as a function of Stoke's radius. The apparent Stoke's radii for CTD_0 and CTDa are 85.5×10^{-8} cm and 37.5×10^{-8} cm, respectively, as indicated by *arrows*.

to the ratio observed in the *in vivo* labeled mouse CTD (see accompanying paper), supporting the idea that CTD kinase *E2* or an enzyme with a similar specificity phosphorylates the CTD *in vivo*. In this paper we show that phosphorylation of a mouse CTD substrate *in vitro* causes a shift in electrophoretic mobility similar to that seen when the CTD is phosphorylated *in vivo*. Furthermore, the CTD adopts a far more extended structure as a result of this phosphorylation.

From analysis of the intermediates in the CTD phosphorylation reaction, we conclude that CTD phosphorylation proceeds by a stochastic, rather than a processive mechanism. The reactions shown in Fig. 1 are carried out in a >40:1 excess of substrate. Thus, if the reaction is processive we should see fully shifted CTD while some of the substrate has not yet been phosphorylated. That this is not so is most clearly evident in Fig. 1*b* where the substrate is prelabeled with casein kinase II. Fully phosphorylated CTD does not appear until all of the substrate is partially phosphorylated. Two recent papers have reported the isolation of protein kinases that shift the mobility of the largest subunit in what appears to be a processive fashion (Lee and Greenleaf, 1989; Payne *et al.*, 1989). These studies used intact RNA polymerase or CTD-fusion protein substrates, and it is possible that in these cases the CTD kinase interacts with the polymerase or fusion protein in such a way that the enzyme is tethered near to the substrate giving rise to apparent processivity. Whether a similar situation arises *in vivo* has not been determined.

The ability to separate different subforms of RNA polymerase II by SDS-PAGE has given rise to the notion that two distinct subspecies exist *in vivo*, the unphosphorylated IIa form and the phosphorylated II_0 form. By examining the time course of phosphorylation shown in Fig. 1*a*, it is evident that SDS-PAGE suffers from severe limitations because the degree of mobility retardation is not proportional to the level of phosphorylation. Thus, when SDS gels are used to analyze II_0 labeled *in vivo*, considerable changes in the level of phosphorylation might go undetected. Indeed, at the later times of the reaction shown in Fig. 1*a* there is up to a 4-fold change in the level of phosphorylation with a minimal change in electrophoretic mobility. The urea-PAGE system described here is more sensitive to changes in CTD phosphorylation and may allow the resolution of multiple II_0 subspecies that are not evident from SDS-PAGE experiments.

Another source of heterogeneity in CTD phosphorylation comes from the observation that only 15–20 of the possible phosphorylation sites are occupied after phosphorylation by CTD kinase *in vitro*. Digestion by proteases has shown that these phosphates are initially added preferentially to the non-consensus carboxyl-terminal repeats, but at the end of the reaction the phosphates are fairly evenly distributed. Whether the same 15–20 sites are used on every CTD molecule, or whether 15–20 sites are selected at random has not been determined.

We have considered several possible explanations for the limited use of possible CTD kinase phosphorylation sites. First, the kinase we have studied may be one of several kinases in mouse cells that phosphorylate the CTD. In fact, we have isolated a second CTD kinase activity from mouse cells (Cisek and Corden, 1990). Furthermore, other kinases have been described in yeast, plant, and human cells that differ in some respects from the mouse CTD kinase (Guilfoyle, 1989; Lee and Greenleaf, 1989; Payne *et al.*, 1989; Stevens and Maupin, 1989). Perhaps different CTD kinases can phosphorylate distinct or partially overlapping sets of sites in the CTD. A second possible reason that not all sites are phosphorylated could be is that the addition of 15–20 phosphates could result in an inhibitory concentration of negative charges on the substrate. Finally, phosphorylation may alter the structure of the CTD such that neighboring unphosphorylated sites are not recognized by CTD kinase. Consistent with this interpretation is our observation that CTD kinase will phosphorylate peptides containing four heptapeptide repeats more readily than those with three. Heptapeptides containing only two repeats do not serve as substrate, presumably because they cannot fold into the appropriate secondary structure.

Although the structure of the CTD is not known, empirical predictive techniques (Chou and Fasman, 1978) suggest a low probability of α-helix or β-sheet formation and a higher probability of turns. Any regularly repeating structure is likely to be extended and have helical symmetry. Models of such structures have been predicted from energy minimization of a polyproline helix (Matsushima *et al.*, 1990). From consideration of the $S_{20,w}$ values and Stoke's radii determined here,

we conclude that the unphosphorylated CTD has a relatively compact structure. Upon phosphorylation this structure unfolds such that the Stoke's radius significantly increases. One possibility is that serine hydroxyl groups are originally involved in hydrogen bonding and that phosphorylation of these hydroxyl groups eliminates these hydrogen bonds and destabilizes the folded structure. Alternatively, the negative charges provided by the phosphates may offer sufficient repulsive force to destabilize the CTD. Turn structures have been proposed for SPXX motif in the CTD (Suzuki, 1990) and to comprise the $p34^{cdc2/CDC28}$ recognition site (Moreno and Nurse, 1990).

Although the function of the CTD is not known, several models have been proposed (Corden *et al.*, 1985; Allison *et al.*, 1985; Sigler, 1988). Perhaps the most interesting possibility is that the CTD is involved in formation of the preinitiation complex of promoter DNA, transcription factors, and RNA polymerase II. Dahmus and colleagues (Laybourn and Dahmus, 1990) have shown that IIa is converted to II_0 after the formation of the preinitiation complex but before transcription initiation. Phosphorylation and the associated conformational change in the CTD could serve to release the CTD from interactions with transcription activating factor-promoter DNA complex. One implication of this model is that CTD kinase is in the preinitiation complex. We have observed that our CTD kinase activity partitions with transcription factors E and F.[3] Further studies of CTD kinase and the preinitiation complex will be necessary to more carefully define the role of CTD phosphorylation.

Acknowledgment—We thank Dr. Lars Cisek for purification of kinases and contributing Fig. 1*b*.

REFERENCES

Ahearn, J. M., Bartolomei, M. S., West, M. L., Cisek, L. J., and Corden, J. L. (1987) *J. Biol. Chem.* **262,** 10695–10705

Allison, L. A., Moyle, M., Shales, M., and Ingles, C. J. (1985) *Cell* **42,** 599–610

Allison, L. A., Wong, J. K-C., Fitzpatrick, V. D., Moyle, M., Shales, M., and Ingles, C. J. (1988) *J. Mol. Cell. Biol.* **8,** 321–329

Bartolomei, M. S., Halden, N. F., Cullen, C. R., and Corden, J. L. (1988) *J. Mol. Cell. Biol.* **8,** 330–339

Cadena, D. L., and Dahmus, M. E. (1987) *J. Biol. Chem.* **262,** 12468–12474

Cantor, C. R., and Schimmel, P. R. (1980) in *Biological Chemistry, Part III*, pp. 1019–1039, Freeman Publications, San Francisco

Chou, P. Y., and Fasman, G. D. (1978) *Annu. Rev. Biochem.* **47,** 251–276

Cisek, L. J., and Corden, J. L. (1989) *Nature* **339,** 679–684

Cisek, L. J., and Corden, J. L. (1990) *Methods Enzymol.*, in press

Corden, J. L. (1990) *Trends Biochem. Sci.* **15,** 383–387

Corden, J. L., Cadena, D. L., Ahearn, J. M., and Dahmus, M. E. (1985) *Proc. Natl. Acad. Sci. U. S. A.* **82,** 7924–7938

Dahmus, M. E. (1981) *J. Biol. Chem.* **256,** 3319–3325

Guilfoyle, T. J. (1989) *Plant Cell* **1,** 827–836

Kim, W.-Y., and Dahmus, M. E. (1986) *J. Biol. Chem.* **261,** 14219–14225

Laemmli, U. K. (1970) *Nature* **227,** 680–685

Laybourn, P. J., and Dahmus, M. E. (1989) *J. Biol. Chem.* **264,** 6693–6698

Laybourn, P. J., and Dahmus, M. E. (1990) *J. Biol. Chem.* **265,** 13165–13173

Lee, J. M., and Greeleaf, A. L. (1989) *Proc. Natl. Acad. Sci U. S. A.* **86,** 3624–3628

Li, W.-B., Bzik, D. J., Gu, H., Tanak, M., Fox, B. A., and Inselburg, J. (1989) *Nucleic Acids Res.* **17,** 9621–9636

Matsushima, N., Creutz, C. E., and Krestinger, R. H. (1990) *Proteins* **7,** 125–155

Moreno, S., and Nurse, P. (1990) *Cell* **61,** 549–551

Morrissey, J. H. (1981) *Anal. Biochem.* **117,** 307–310

Nonet, M., Sweetser, D., and Young, R. A. (1987) *Cell* **50,** 909–915

Nurse, P. (1990) *Nature* **334,** 503–508

Payne, J. M., Laybourn, P. J., and Dahmus, M. E. (1989) *J. Biol. Chem.* **264,** 19621–19629

Sigler, P. B. (1988) *Nature* **333,** 210–212

Stevens, A., and Maupin, M. K. (1989) *Biochem. Biophys. Res. Commun.* **159,** 508–515

Suzuki, M. (1990) *Nature* **344,** 562–565

Woychik, H., and Young, R. A. (1990) *Trends Biochem. Sci.* **15,** 347–351

Zandomeni, R., Zandomeni, M. C., Shugar, D., and Weinmann, R. (1986) *J. Biol. Chem.* **261,** 3414–3419

Zehring, W. A., Lee, J. M., Weeks, J. R., Jokerst, R. S., and Greenleaf, A. C. (1988) *Proc. Natl. Acad. Sci. U. S. A.* **85,** 3698–3702

[3] E. Barron and J. Corden, unpublished results.